Wi-Gym: Gymnastics Activity Assessment Using Commodity Wi-Fi

Lei Zhang, *Member, IEEE*, Wenyuan Huang, Xiaoxia Jia, Xiaojie Fan, Xiaochen Fan, Liangyi Gong, *Member, IEEE*, Wenyuan Tao, and Shiwen Mao, *Fellow, IEEE*

***Abstract*—Practicing gymnastics activities at home with online resources has become an increasingly popular choice due to its convenience and accessibility. However, without face-to-face guidance by a trainer, a major challenge is how to assess the quality of performed gymnastics activities, effectively and fairly. Existing intrusive assessing approaches usually require live cameras or wearable sensors, which usually generate privacy and feasibility concerns. There is a lacking of accurate approaches to assess the quality of the activities. To address these challenges, a gymnastics activity assessment approach is proposed in this article, and Wi-Gym, an effective first-of-its-kind gymnastics activity assessment system is developed utilizing commodity Wi-Fi. Wi-Gym is designed to compare the activity-induced channel state information (CSI) dynamics by an exerciser and that of a trainer utilizing dynamic time warping (DTW). The comparison results are provided by a fuzzy inference system (FIS). To make Wi-Gym robust to the changes in the environment, domain adaptation is leveraged to mitigate the data distribution imbalance caused by the environment changes. Extensive experimental studies have been conducted using Wi-Gym, acoustic, and video-based sensing systems. The experimental results validate the effectiveness and robustness of the proposed approach.**

***Index Terms*—Activity recognition, channel state information (CSI), gymnastics activity assessment, Wi-Fi sensing.**

Manuscript received 21 April 2022; accepted 1 July 2022. Date of publication 6 July 2022; date of current version 21 November 2022. This work was supported by NSFC under Grant 61772364 and Grant 61902211. The work of Shiwen Mao was supported in part by NSF under Grant ECCS-1923163. *(Wenyuan Huang and Xiaoxia Jia are co-second authors with equal contributions.) (Corresponding authors: Wenyuan Tao; Shiwen Mao.)*

Lei Zhang is with the College of Intelligence and Computing and the Tianjin Key Laboratory of Advanced Network Technology and Application, Tianjin University, Tianjin 300050, China, and also with the Key Laboratory of Grain Information Processing and Control, Henan University of Technology, Ministry of Education, Zhengzhou 450001, Henan, China (e-mail: lzhang@tju.edu.cn).

Wenyuan Huang, Xiaoxia Jia, and Xiaojie Fan are with the College of Intelligence and Computing and the Tianjin Key Laboratory of Advanced Network Technology and Application, Tianjin University, Tianjin 300050, China (e-mail: lzhang@tju.edu.cn; wy_huang1998@tju.edu.cn; jiaxiaoxia@tju.edu.cn; funfunfunny@tju.edu.cn).

Xiaochen Fan is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Consulting Center for Strategic Assessment, AMS, Beijing 100091, China (e-mail: fanxiaochen33@gmail.com).

Liangyi Gong is with the School of Software and BNRist, Tsinghua University, Beijing 100084, China (e-mail: gongliangyi@gmail.com).

Wenyuan Tao is with the College of Intelligence and Computing, Tianjin University, Tianjin 300050, China (e-mail: taowenyuan@tju.edu.cn).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Digital Object Identifier 10.1109/JIOT.2022.3188916

I. Introduction

GYMNASTICS can be practiced at a playground, a gym, or a sports field, where a trainer or instructor is usually involved to guide the exercisers. Under the recent global pandemic of COVID-19, there is a significant shift of gymnastics practices to home scenarios. Typically, people can perform gymnastics activities at home by following an online tutorial, as illustrated in Fig.1. Despite that the "home-gym" has the advantages of convenience and accessibility, it becomes questionable whether the exercisers really follow the standard baselines and whether they can improve their health as well as prevent possible physical injuries.

To assess the quality of home-gym exercises, both subjective and objective approaches have been proposed in the past. For objective assessment, direct evaluation and guidance to exercisers are done by visual inspection or using a video camera. But this approach has several drawbacks in terms of privacy violation, lighting condition constraints, and security concerns [1]. These disadvantages prevent the wide adoption of vision-based assessment systems for home gymnastics. Wearable devices are used in objective gymnastics evaluation systems [2]–[9]. However, the high complexity of gymnastic exercises makes it difficult to precisely evaluate the specific activity's quality. Moreover, it would be inconvenient, unsafe, and infeasible for exercisers to always wear the devices while doing gymnastics exercises. To this end, the ubiquitous and noninvasive radio frequency (RF) sensing becomes a promising candidate for home gymnastics assessment.

The pervasive deployment of Wi-Fi [10], [11] has triggered extensive research on Wi-Fi-based human activity sensing. The Wi-Fi channel state information (CSI), which offers fine-grained granularity information about the propagation environment [12], provides opportunities to detect human presence and identify their activities through a series of data processing techniques. Various Wi-Fi-based human activity recognition systems have been proposed, such as Wi-Fall [13], RT-Fall [14], and FarSense [15]–[17]. However, these systems focus on detecting an instant movement that generates explicitly observable signal variations. There are hardly any prior work on assessing the quality of gymnastics activities (e.g., whether the movements of exercisers follow the standard baselines) [18].

There are many challenges for Wi-Fi-based gymnastics exercise evaluation system. First, a gymnastics exercise consists of a series of activities and each activity involves the

Fig. 1. Scenario of a home-gym.

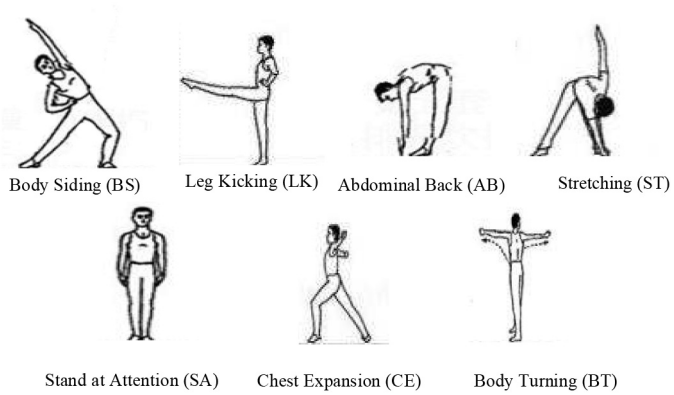


Fig. 2. Seven basic gymnastics activities in our experiments.

movements of multiple body parts. In addition, different people perform each activity differently, while during the gymnastics exercise, different body parts move at different speeds for a specific activity. The Wi-Fi signals reflected by different body parts are mixed with static multipath signals at the Wi-Fi receiver. Due to the lack of spatial information in CSI, it is challenging to segment the entire continuous gymnastics exercise into a series of individual activities, and then identify and assess each individual activity. Second, due to the lack of spatial resolution, RF signals, as well as the features, highly depend on the deployment environment where the gymnastic activities are performed. The primitive RF signals usually carry adverse environment information irrelevant to the activities. Therefore, the CSI variations caused by the environmental changes may lead to incorrect activity recognition. Therefore, the classifier trained in one scenario usually suffers drastic performance degradation in another scenario. When there is an environmental change, considerable efforts in either data collection or model retraining are required to maintain a satisfactory accuracy. It is challenging to reduce such system maintenance cost. Third, how to effectively and fairly assess each activity is challenging. An effective strategy, including assessment reference and baseline standard, needs to be developed to evaluate each activity.

As one kind of typical gymnastics exercise, the so-called "radio gymnastics" is very popular in China. As illustrated in Fig. 2, it mainly consists of seven basic nonrepetitive activities. Using the radio gymnastics exercise as an example, we propose a ubiquitous and nonintrusive gymnastics exercise assessment system, called Wi-Gym in this article. Wi-Gym is designed to identify and rate the gymnastics activities. In practice, Wi-Gym can be customized to rate various types of exercise activities according to the specific application requirements. The main contributions of this article are summarized as follows.

1) A complete set of gymnastics exercise is segmented into a number of individual activities. Activity recognition and rating are performed for each segmented activity. Specifically, fast Fourier transform (FFT) is employed to profile the power distribution of CSI phases. An energy indicator is introduced to capture the significant changes in CSI phase difference for automatic activity segmentation. A 2-D Gabor filter is then leveraged to extract features automatically and a classification model based on support vector machine (SVM) is adopted to recognize each segmented gymnastics activity.
2) To relief the system maintenance efforts, domain adaptation is introduced to mitigate the impact of imbalanced data distribution. Specifically, the classifier is trained with labeled data and knowledge is derived for known scenarios. With the derived knowledge, the classifier model, when deployed in a new scenario, can be trained with only a few labeled data and maintain a fairly well performance. This effective across-domain knowledge reuse approach mitigates the impact of different environments and maintains a consistently satisfactory performance.
3) In order to rate the quality of each individual activity, an effective strategy is proposed. Specifically, the trainer activity-induced CSI data are regarded as references. The similarity between the activity performed by an exerciser and by the trainer is quantified and measured. Specifically, the feature vectors derived from the exerciser activity-induced CSI dynamics and the trainer activity-induced CSI dynamics are compared using the dynamic time warping (DTW) method. Assessment is provided by a fuzzy inference system (FIS) based on the comparison results.
4) We implement a prototype of the Wi-Gym using off-the-shelf Wi-Fi devices and conduct extensive experiments in real-world scenarios. Overall, we test the performance of Wi-Gym on the seven basic radio gymnastics activities as shown in Fig. 2. For comparison purposes, the proposed approach is also implemented and tested using two other sensing methods, i.e., acoustic and video camera. We find all the three prototype systems can achieve a satisfactory performance. The experimental results validate the effectiveness of the proposed gymnastic activity assessment approach. We believe that this article can open a new direction of RF sensing for smart home and e-health.

The remainder of this article is organized as follows. In Section II, we review the prior works on RF sensing-based activity recognition. In Section III, we present the detailed

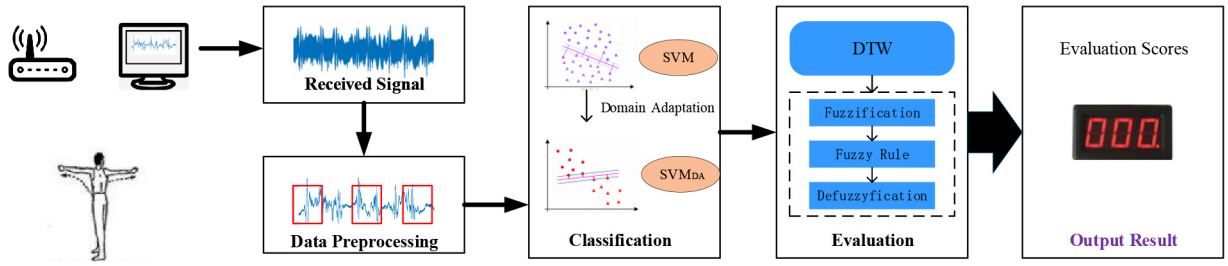


Fig. 3. Architecture of the proposed Wi-Gym system.

design and implementation of Wi-Gym. Then, we analyze the experimental results in Section IV and discuss the limitations of Wi-Gym in Section V. Finally, we conclude this article in Section VI.

II. Related Work

A. Commodity Wi-Fi-Based Activity Recognition

With the wide application and in-depth research of Wi-Fi, considerable advances have been achieved in the field of activity perception and recognition based on Wi-Fi. The research of Wi-Fi-based activity recognition can be roughly divided into two categories: 1) learning-based approaches [12], [14], [19]–[22] and 2) nonlearning-based approaches [23]–[27]. For example, E-eyes [19] is a typical learning-based activity recognition system, which uses the feature profile of signals for activity recognition. In WiKey [28], the typed keys are recognized by collecting extensive data to train the models. RT-Fall [14] is proposed for fall detection by learning the pattern of falls in the time–frequency domain. WiDir [23] is based on the Fresnel zone model to detect the walking direction by analyzing the Wi-Fi phase difference data. In CARM [12], the patterns of movement speeds of different activities are used to build the corresponding hidden Markov model (HMM). Although commodity Wi-Fi-based sensing is proven effective for activity recognition, there is hardly any prior work on nonrepetitive activity evaluation.

B. Activity Evaluation

Research on activity evaluation is also scarce, and the majority of work has been done during the last decade. For instance, in [3] and [4], a human body sensor network is utilized to provide inertial data for baseball swing evaluation. The swing motion is recorded, the intersegment coordination is measured, and the user feedback is generated. The basic principle of sensor-based measurement is that simultaneous interpreting of signals from different sensors can result from different human activities. In [29], an optical motion capturing system based on the mark is proposed, which constitutes the capturing component of a dance training system. The data obtained according to joint position, velocity, and angle are matched with the data recorded in advance. The visual as well as numerical feedback are subsequently provided. The marker-based optical approach cannot work without an appropriate lighting condition. Since Wi-Fi is ubiquitous and nonintrusive, a Wi-Fi-based solution will be a promising candidate. So far, there is hardly any prior work employing commercial off-the-shelf (COTS) Wi-Fi to evaluate the nonrepetitive gymnastics activities. To the best of our knowledge, Wi-Gym is the first commodity Wi-Fi-based nonrepetitive gymnastics exercise evaluation system.

III. Wi-Gym Design

A. System Architecture

The key ideas of Wi-Gym are to accurately recognize and assess the gymnastics activities, as well as mitigating the influence of different deployment environments. As shown in Fig. 3, the proposed Wi-Gym consists of data collection and preprocessing, activity segmentation and recognition, and assessment of each individual activity. Specifically, CSI data are collected and calibrated by the signal collection module. The continuous nonrepetitive gymnastics activity-induced CSI signals are then partitioned by the activity segmentation module for each individual activity. In the activity recognition module, an SVM-based domain adaptation approach is proposed to achieve effective across-environment knowledge reuse, mitigate the environmental influence, and maintain a satisfactory activity identification accuracy. In order to assess the quality of an individual gymnastics activity, the gymnastics activity performed by a trainer is adopted as a benchmark. The similarity between an exerciser's CSI pattern and the trainer's CSI pattern recorded beforehand are evaluated using the DTW method. The final evaluation scores are provided by the FIS based on the results of DTW.

B. Noise Removal

Usually, Wi-Fi CSI values are too noisy to be directly used for human activity recognition. Even in a static environment without any human activity, CSI values fluctuate because Wi-Fi devices are susceptible to surrounding electromagnetic noises. Moreover, the internal state changes in Wi-Fi devices, e.g., transmission rate adaptation and transmission power adaptation often introduce impulse and burst noises in CSI values. In this article, we propose a two-level denoising method that applies the Hampel identifier and principal component analysis (PCA) to retain the useful information in CSI. First, CSI dynamics that are not caused by human activity are removed by using the Hampel identifier. The Hampel identifier calculates the median and the standard deviation of the sliding window consisting of the current amplitude and its surrounding amplitudes. If the absolute value of the difference between the current amplitude value and the median is

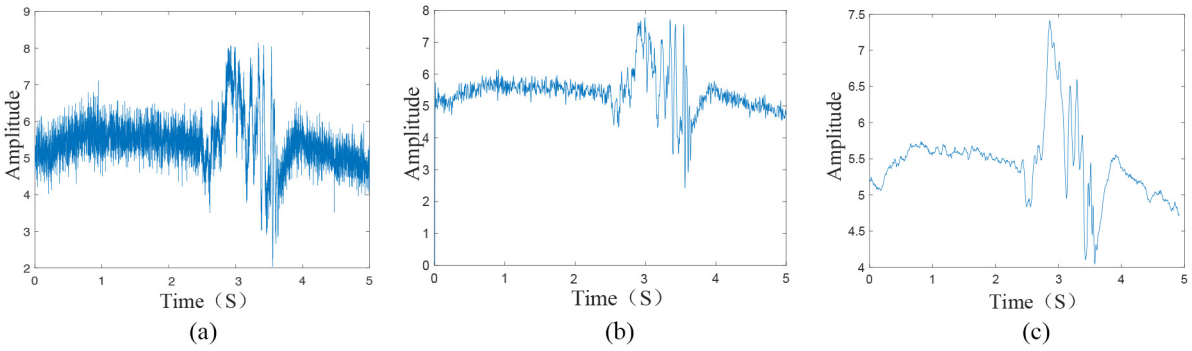


Fig. 4. Denoised Wi-Fi CSI data using the two-level calibration method. (a) Raw CSI. (b) CSI after Hampel filter. (c) CSI after PCA.

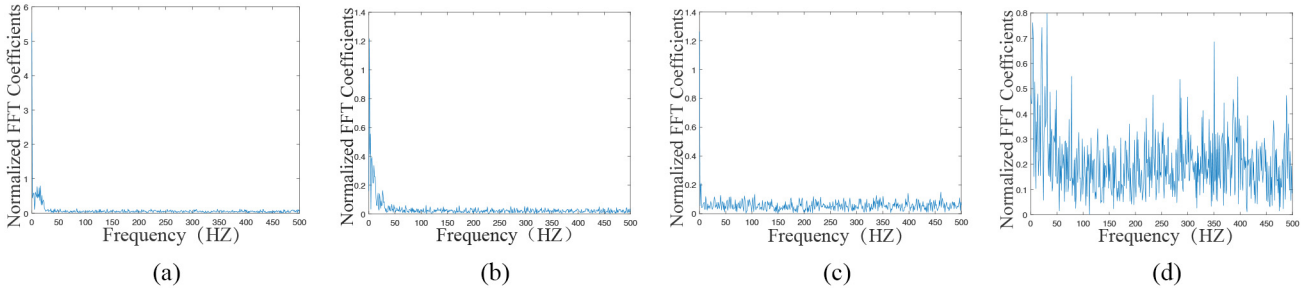


Fig. 5. Power distribution of FFT profiles in four different scenarios. (a) Empty room. (b) Subject walking into the room. (c) Subject does no activity. (d) Subject performs a gymnastics exercise routine.

more than a threshold (i.e., a given number times the standard deviation), the current amplitude value will be identified as an outlier. The outlier can be effectively removed by replacing it with the median. Second, a CSI denoising scheme based on PCA is proposed. This scheme is based on the fact that the signal fluctuations caused by body motion in all subcarriers of the observed CSI values are correlated. In short, PCA has two advantages. First, the PCA can reduce the computational complexity because it reduces the dimensionality of the signals acquired from the 30 subcarriers in each TX–RX pair, while retaining valuable information about the activity. Second, PCA can eliminate the in-band noise that traditional low-pass filters cannot. This takes advantage that CSI variations caused by human movements are correlated over all subcarriers, while noisy components are usually not correlated. Fig. 4 illustrates the denoised CSI stream using the two-level denoising method.

C. Activity Segmentation

In Fig. 5, the power distribution of an empty room, when there is one stationary subject in the room, when there is one subject walking in the room, and when there is one subject performing a gymnastics exercise in the room are presented. From Fig. 5, we can see that there is a significant drop in power from high frequency to low frequency between the case of doing gymnastic exercise and the case of no activity at all. In addition, there is always a static state between the two sequential activities during the gymnastic exercise routine. Therefore, the state switching generates power distribution fluctuations. This coincides with that in Fig. 5(d).

In the absence of an activity, there is only a smaller variance caused by background noise. The noise level can hardly change because of the stable environment. A dynamic threshold algorithm is employed to check the noise level. An exponential moving average algorithm is utilized for updating the noise level $L(t)$ during the silent period, given by

$$L_t = (1-\tau)L_{t-1} + \tau \times \mathrm{VAR}_t \tag{1}$$

where the coefficient τ is set to 0.05 empirically, and the variance VAR_t for the ith sliding window with 200 samples is calculated in the second principal component. If VAR_t is larger than four times of the noise level in one sliding window, the starting point of the next activity will be detected. Similarly, the end point of the activity can also be discovered. Thus, the segmentation of the continuous gymnastics routine is done this way.

In Wi-Gym, the optimal sliding window size is set to 3 s. Two seconds for the activity duration and 1 s for the pause period. FFT frequency analysis is utilized to capture the patterns of gymnastics activities and produce the spectra.

D. Activity Recognition

Each gymnastics exercise consists of a series of activities, while each activity involves in the movements of multiple body parts. Different people usually perform the same activity differently. In addition, during the gymnastics exercise, different body parts move at different speeds for each specific activity and Wi-Fi signals reflected by different body parts are mixed with the static multipath components at the Wi-Fi receiver. Therefore, feature extraction is nontrivial due to the nonlinear signal variations in the CSI time series. Due to the limitations, the quality of extracted features cannot be guaranteed by traditional feature extraction approaches.

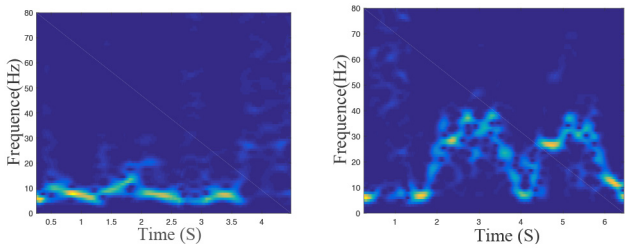


Fig. 6. Spectrograms of two activities (i.e., squatting and stagnation) conducted by the same exerciser in the same environment. (a) Squat. (b) Walk.

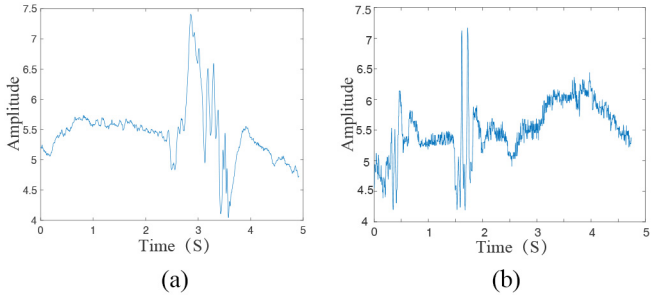


Fig. 7. CSI time-series of doing the same activity in two different environments. (a) Office. (b) Laboratory.

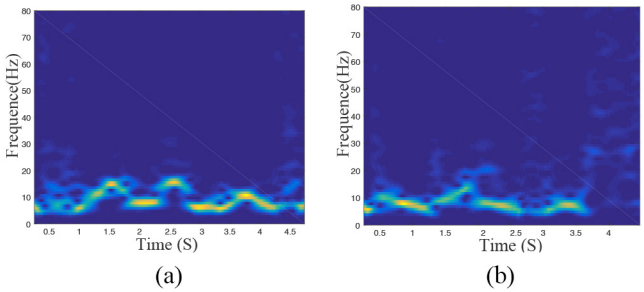


Fig. 8. Spectrograms of doing the same activity in two different scenarios. (a) Office. (b) Laboratory.

Hence, an automatic feature extraction method called Gabor filter [30] is leveraged in this article. The advantages of the Gabor algorithm are as follows: 1) the feature extraction does not depend on experience; 2) the feature extraction is not influenced by change of scenario; and 3) the number of features obtained can be large. The shape of the human body is complicated, and different body parts move at various speeds. Therefore, the different activities' frequency distributions derived from the corresponding spectrograms have distinct textures, as shown in Fig. 6. In addition, the Gabor filter can extract a large number of abstract features, which can address the shortcomings of traditional feature extraction methods.

The Gabor filter is defined as follows:

$$G(i, j, \psi, \sigma, \gamma, \lambda, \theta) = \exp\left\{2\pi \frac{i'}{\lambda} + \psi\right\} \exp\left\{-\frac{i'^2 + \gamma^2 j'^2}{2\sigma^2}\right\} \tag{2}$$

where (i, j) is the coordinate position in the picture (i.e., the spectrogram), and phase ψ, aspect ratio σ, bandwidth γ, wavelength λ, and rotation angle θ are the five parameters to define the Gabor filter. ln (2), (i', j') is computed as follows:

$$\begin{cases} i' = i\cos\theta + j\sin\theta \\ j' = -i\sin\theta + j\cos\theta. \end{cases} \tag{3}$$

The real part and the imaginary part of (2) are, respectively, given by

$$G_{\text{real}}(i, j, \psi, \sigma, \gamma, \lambda, \theta) = \exp\left\{-\frac{i'^2 + \gamma^2 j'^2}{2\sigma^2}\right\} \cos\left(2\pi \frac{i'}{\lambda} + \psi\right)$$

$$G_{\text{imag}}(i, j, \psi, \sigma, \gamma, \lambda, \theta) = \exp\left\{-\frac{i'^2 + \gamma^2 j'^2}{2\sigma^2}\right\} \sin\left(2\pi \frac{i'}{\lambda} + \psi\right).$$

The direction of the Gabor filter is set to 5, the size is set to 39, the scale is set to 8, and the vertical and horizontal downsampling intervals are set to 4 and 6, respectively, in Wi-Gym. The input spectrogram is reconstructed into a matrix form with size 30×200. The dimension of the extracted features is about 10 880. Finally, the extracted features are fed into the classifier for recognition.

E. SVM-Based Domain Adaptation

In practice, a gymnastics activity-induced CSI dynamics in one scenario is different from that in another scenario, as illustrated in Fig. 7(a) and (b). This is due to the environmental dependence and has a considerable negative impact on activity identification. However, as shown in Fig. 8(a) and (b), the corresponding trends of frequency change caused by doing the same activity in different environments are similar. This implies the correlation of the scenarios and provides a solid foundation for domain adaptation applied to cross-scenario gymnastic activity identification.

The traditional SVM model works well when the training data and test data share the same distribution [31]. When the environment changes, the data distribution will very likely be different, resulting in poor performance of the SVM model. There have been many prior works studying the domain adaptation problem, which aims to balance the data distribution in different environments. Among them, the approaches based on the SVM classifier are the most popular ones [32], [33], largely due to their high efficiency and low overhead. To mitigate the negative impacts caused by changes in the environment, a domain adaptation model based on the SVM classifier, called SVM-DA, is proposed.

Specifically, the activity-induced CSI dynamics collected in one scenario is regarded as the source domain. The same activity-induced CSI dynamics collected in a different scenario is regarded as the target domain. When the SVM classifier is trained with a small amount of target domain data, its performance will be poor [31]. This is improved by the following domain adaptation methodology. The trained parameters w_s in the source domain are brought to the target domain to further optimize the SVM model as

$$\min_{w_t, b_t, \varphi} \quad P(w_t, \varphi) = \eta \|w_t - w_s\|^2 + \frac{1}{2} w_t^T w_t + D \sum_{i=1}^{M} \varphi_i \tag{4}$$

$$\text{s.t.} \quad l_i\left[\left(w_t^T * f_i\right) + b_t\right] \geq 1 - \varphi_i, \quad i = 1, 2, \ldots, M \tag{5}$$

$$\varphi_i \geq 0, \quad i = 1, 2, \ldots, M \tag{6}$$

where D and η are the corresponding penalty parameters; w_s are the weights of the source domain and w_t are the weights of the target domains; f_i is the feature vector marked by l_i, which is the ith feature among the M extracted features; and φ_i is a slack variable. In (4), the $(1/2)w_t^T w_t + D\sum_{i=1}^{N}\varphi_i$ term is the objective function of the traditional SVM model. By adding the term $\eta\|w_t - w_s\|^2$ in the objective function of SVM-DA, the source domain learning is introduced to represent the variance of w_t and w_s classifiers.

The Lagrangian of problem (4) is given by

$$\mathcal{L}(w_t, b_t, \varphi; \alpha, \beta) = P(w_t, \varphi) - \sum_{i=1}^{M}\alpha_i\left\{l_i\left[\left(w_t^T * f_i\right) + b_t\right] - 1 + \varphi_i\right\} - \sum_{i=1}^{M}\beta_i\varphi_i \tag{7}$$

where $\alpha_i \geq 0$, $\beta_i \geq 0$, $i = 1, 2, \ldots, M$, are the Lagrange multipliers. According to [34], when the Karush–Kuhn–Tucker (KKT) condition in [35] is satisfied, the dual problem can have the same optimal solution as the original problem. This problem is solved by replacing w_t in the Lagrange function as

$$\max_{\alpha} \; S(\alpha) = -\frac{1}{2(2\eta+1)}\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j\beta_i\beta_j G(f_i, f_j) - \sum_{i=1}^{M}\alpha_i\left(\frac{2\eta w_s l_i f_i}{2\eta+1} - 1\right) + \frac{\eta}{2\eta+1}w_s^T w_s \tag{8}$$

$$\text{s.t.} \; \sum_{i=1}^{M}\alpha_i l_i = 0, \; 0 \leq \alpha_i \leq D, \quad i = 1, 2, \ldots, M \tag{9}$$

where $G(f_i, f_j)$ is the Gaussian kernel, which is used to map the input samples to the high-dimensional feature space. Then, the sequential minimum optimization (SMO) algorithm proposed in [36] is used to solve the above problem, and the optimal solution is α_i^*, $i = 1, 2, \ldots, M$, and

$$w_t^* = \frac{\sum_{i=1}^{M}\alpha_i^* l_i f_i + 2\eta w_s}{2\eta+1} \tag{10}$$

where $(2\eta w_s/2\eta+1)$ is the knowledge learned from the source domain, and $([\sum_{i=1}^{M}\alpha_i^* l_i f_i]/2\eta+1)$ is the new knowledge learned from the target domain. Then, b_t^* can be calculated by

$$b_t^* = \frac{1}{|K|}\sum_{k\in K}\left(l_k - \frac{2\eta w_s f_k + \sum_{i=1}^{K}\alpha_i^* l_i G(f_i, f_k)}{2\eta+1}\right) \tag{11}$$

where (f_k, l_k) represents any support vector and K is a set containing all the support vectors. For the target domain test sample f_t, the final discriminant function can be calculated as follows:

$$\mathcal{L}(f) = \text{sign}\left(\frac{2\eta w_s f_t + \sum_{i=1}^{M}\alpha_i^* l_i G(f_i, f_t)}{2\eta+1} + b_t^*\right). \tag{12}$$

Therefore, the sign$(\cdot)$ function can be leveraged to label the test sample f_t. This verifies the feasibility of SVM-DA.

The most significant task of Wi-Gym is to rate each gymnastic activity. This is achieved by a comparison of the activity performed by an exerciser and the same activity performed by a trainer. Each activity is portrayed by the extracted features. DTW is adopted to calculate the distance between the two feature vectors. By exploiting the idea of dynamic programming, DTW measures the distance between two feature vectors on the premise of resisting signal distortion. Consider two feature vectors: $F = (f_1, f_2, \ldots, f_m)$ of length m and $V = (v_1, v_2, \ldots, v_n)$ of length n, where $m \neq n$. The distance between the vectors is calculated as

$$\text{Dis}(F, V) = \sum_{i=1}^{m}\sum_{j=1}^{n}\text{DTW}(f_i, v_j). \tag{13}$$

The final scores are provided by the FIS [37] based on the DTW results, which is described as follows.

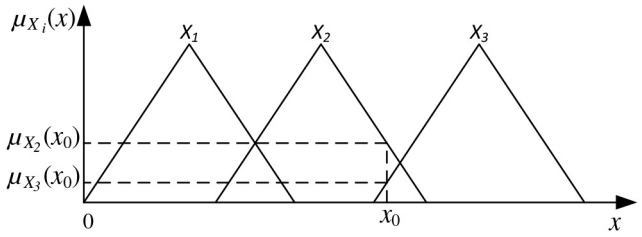


Fig. 9. Fuzzy set system with triangular membership function input.

F. Fuzzy Inference System

The comparison results of two activities are multifactor, comprehensive, and uncertain. The FIS is leveraged to ensure the fairness of rating. Specifically, each relevant factor is assigned with a judging index according to the given conditions. The general judgment can then be obtained. The main components of the FIS include 1) fuzzification, which establishes the fuzzy relation of input data and fuzzy membership function; 2) fuzzy inference, which determines the fuzzy logic rules; and 3) defuzzification, where the fuzzy output is converted into a score of activity evaluation. The detailed implementation of each component is discussed as follows.

1) Fuzzification: The DTW of the feature vectors is the input to fuzzy inference. Fuzzification is the process of transforming the input values into the element of each set. The triangle membership function is adopted to determine the relationship between the input value and the membership degree through the verification of a large number of experimental data. As shown in Fig. 9, the triangle membership function is given by

$$\mu_{K_i}(x) = \begin{cases} 0, & x \leq m \\ (x-m)/(n-m), & m < x \leq n \\ (t-x)/(t-n), & n \leq x < t \\ 0, & x \geq t \end{cases} \tag{14}$$

where parameters m and n determine the foot of the triangle and parameter t determines the peak of the triangle.

Fig. 9 shows the fuzzy set system with triangular membership function input. For a certain value x_0 of the basic attribute, the membership degree on the fuzzy subset X_i is fuzzified to $\mu_{X_i}(x_0)$. Any particular input is interpreted from this fuzzy set, and a degree of membership is then obtained.

Algorithm 1: Fuzzy Inference Algorithm

Input : Feature vectors of an exerciser's activity $\vec{F} = \{\vec{f_1}, \vec{f_2}, ..., f_x\}$ and feature vectors of a professional's activity $\vec{\xi} = \{\vec{\xi_1}, \vec{\xi_2}, ..., \xi_y\}$;
Output: *EvaluationScore* of the activity

```
K' is the result set , k is the output index and K is the domain of the input set;
D⃗ ← CalculateDist(F⃗, ξ⃗);
for i ← 1 to length (K) do
    μ⃗ ← MembershipTrans(μ_{K_i}(D⃗));
end
μ_{K'}(k) ← MembershipTrans(μ⃗);
return Defuzzification(k, μ_{K'}(k));

function CalculateDist(F⃗, ξ⃗):
    Initialize D⃗;
    D⃗ ← DTW(f⃗_t, ξ⃗);
    return D⃗;
end function

function μ_{K_i}(D⃗):
    According to K_i, initialize foot_1, foot_2, and peak;
    Initialize M⃗ with m_i;
    for i ← 0 to length (D⃗) do
        if d_i < foot_1 then
            m_i ← 0;
        end
        else if d_i < foot_2 then
            m_i ← (d_i − foot_1)/(foot_2 − foot_1);
        end
        else if d_i < peak then
            m_i ← (peak − d_i)/(peak − foot_2);
        end
        else
            m_i ← 0;
        end
    end
    return M⃗;
end function

function MembershipTrans(M⃗):
    maxMembership ← 0;
    for i ← 0 to length (M⃗) do
        if m_i > maxMembership then
            maxMembership ← m_i;
        end
    end
    return maxMembership;
end function

function Defuzzification(m, n):
    return EvaluationScore = ∫mndm / ∫ndm;
end function
```

2) Fuzzy Rule Determination: All membership functions and weights are manually/intuitively adjusted in most FISs. The set of rules for building the fuzzy inference are presented in Algorithm 1.

3) Defuzzyfication: The fuzzy output is transformed into a clear score by the membership value. Fuzzification and defuzzification are reciprocal processes. With the discrete triangular linear functions, the center of gravity (COG) method [38] is adopted as follows:

$$k_0 = \frac{\int k\mu_{K'}(k)dk}{\int \mu_{K'}(k)dk} \tag{15}$$

where $\mu_{K'}(k)$ is the membership function of K', and k_0 is the result of defuzzification, which indicates the rating of the gymnastics activity.

IV. Experiments and Evaluation

A. Prototyping and Experiment Configuration

1) Wi-Gym Prototyping: The proposed Wi-Gym system is implemented using three sensing technologies, Wi-Fi, acoustic, and video camera, the latter two are used for comparison purposes.

1) *Wi-Fi:* Wi-Gym is implemented with two T-series Lenovo laptops. Each laptop is equipped with Intel 5300 network interface card and the Ubuntu desktop 11.04 operating system. One of the laptops is connected with three external antennas as the transmitter and the other is connected with three external antennas as the receiver. The two laptops work in the monitor mode in the 5-GHz band. The sampling frequency used in this article is 1024 Hz.
2) *Acoustic:* In the experiment, a commercial off-the-shelf (COTS) speaker (JBL Jembe, 6 W, 80 dB) works as the transmitter, and a microphone (SAMSON MeteorMic, 16 bit, 48 kHz) works as the receiver. The transmitter and receiver are connected to a Lenovo laptop (Intel Core i5-10210U CPU, 8G RAM). They are placed at a height of 80 cm. The speaker is programmed to send continuous acoustic signals at a constant frequency of 20 kHz and the microphone receives the acoustic signals with a sampling rate of 48 kHz. Then, the microphone sends the captured signals to the connected laptop for data processing. The captured acoustic data go through similar processing stages as the Wi-Fi data, including noise removal, activity segmentation, Gabor filter, SVM with domain adaptation, DTW, and FIS.
3) *Video Camera:* The vision data used for training and testing are collected using an Xbox Kinect 2.0 device. The device is deployed in the LoS scenario at a height of 80 cm. The frame rate of Kinect 2.0 is 30 frames/s. The vision data are manually clipped, such that each resulting video clip contains one of the seven basic activities of the radio gymnastic routine. OpenPose [39] is an open-source multistage CNN-based 2-D image pose detection system. OpenPose generates 2-D key points based on the input images. For better evaluation of the Wi-Gym performance, both the gymnastic expert's and the test subjects' seven basic activities are recorded and processed. They are fed into OpenPose, respectively, which will generate a 2-D skeleton for each subject and gymnastic expert's activity. The 2-D key points go through the stages of SVM-DA, DTW, and FIS. The scores for their activities are finally obtained. The Intel Xeon E5-2650 v4 CPU and NVIDIA Tesla P100 16-GB GPU are

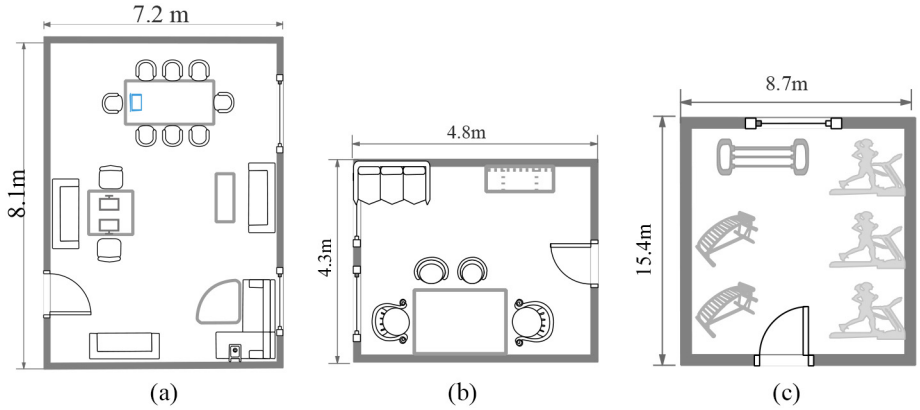


Fig. 10. Layouts of three home-gym scenarios considered in our experimental evaluation. (a) Laboratory. (b) Office. (c) Gymnasium.

utilized for clipped video data, training the CNN model, and testing with the trained model.

2) Experiment Setting: Wi-Gym is evaluated in three typical home-gym environments, including: 1) office; 2) laboratory; and 3) gymnasium, as shown in Fig. 10. The laboratory is 8.1 m × 7.2 m with two tables, ten chairs, and two computers. The office is 4.8 m × 4.3 m with four chairs, one table, and one bookcase. The experiment is also conducted in a 15.4 m × 8.7 m gymnasium, which has three treadmills, two roman chairs, and one barbell.

Seven typical continuous nonrepetitive gymnastic activities are tested. The order of the activities is predetermined according to the radio gymnastics routine. As shown in Fig. 2, these include stretching (ST), chest expansion (CE), leg kicking (LK), body siding movement (BS), body turning movement (BT), standing at attention (SA), and abdominal back movement (AB). Fifty volunteers, including 25 males and 25 females are recruited and tested. They differ in age (18–40), weight (45–90 kg), and height (155–188 cm). Five gymnastic trainers are invited to do the seven basic activities sequentially. Each activity is repeated for 100 times. During the experiment, they wear daily clothes and stand in the LoS path 1 m away. The Wi-Fi, acoustic, and video data are collected simultaneously. For each sensing method, each subject repeats all the seven activities 100 times in each scenario. For each sensing method, totally 105 000 traces are collected, including 40 000 traces from the laboratory, 30 000 traces from the office, and 35 000 traces from the gym. A data set is created involving all the traces with various activity qualities. Most of the experiments are conducted at daytime. There is natural light coming from the windows and no electric light is turned on during the experiment. The ground truth is the average scores provided by the gymnastic coaches. Each subject faces the LoS path and is 1 m away from the LoS path. The final score is averaged over 100 randomly selected scores.

3) Performance Metrics: To analyze the performance of Wi-Gym, true positive rate (TPR) and false positive rate (FPR) are utilized for gymnastics activities assessment. TPR is defined as the rate that "doing gymnastics exercise well" (i.e., the average score is higher than 80) is correctly identified, given by

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{16}$$

and FPR is defined as the rate of misidentification, given by

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN} + \mathrm{FP}} \tag{17}$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively. The accuracy is defined as the ratio of the number of correctly identified instances to the number of total instances.

TABLE I
ASSESSMENT ACCURACY ACHIEVED BY THE WI-FI-BASED PROTOTYPE IN THREE TYPICAL INDOOR ENVIRONMENTS

Scenario	True Positive Rate (TPR)	False Positive Rate (FPR)
Laboratory	89.5%	10.5%
Office	88.9%	11.1%
Gymnasium	86.4%	13.6%

B. Overall Performance

The overall experimental results achieved by the Wi-Fi-based prototype are summarized in Table I. It is obvious that Wi-Gym achieves a satisfactory performance in all the three scenarios, with a mean TPR of 88.27% and a mean FPR of 11.73%. The results show that Wi-Gym exhibits desirable advantages in different scenarios, but there are still some differences. Wi-Gym performs the worst in the gym scenario, with a TPR of 86.4% and an FPR of 13.6%. This is because the layout of the gym is more complex than the other two, and thus the multipath effect is the strongest. In the laboratory scenario, Wi-Gym achieves the best performance, because the layout is the simplest and the multipath effect is not serious. In addition, the receiver operating characteristic (ROC) curve exhibits the tradeoff between TPR and FPR over various settings, as shown in Fig. 11. In practice, Wi-Gym achieves an average TPR and FPR of 93.49% and 7.48%, respectively, at the equal error rate (EER) point. These results validate the feasibility of Wi-Gym for gymnastic activity evaluation.

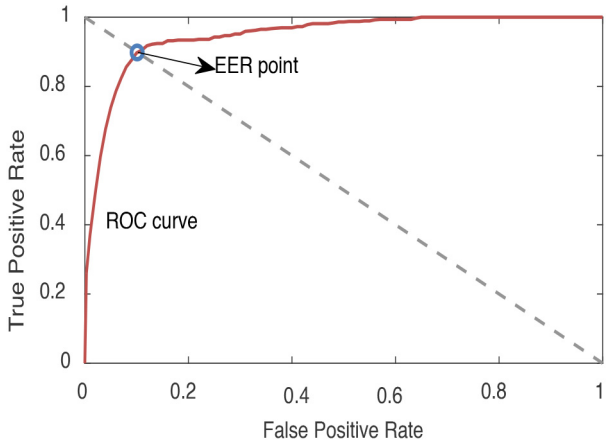


Fig. 11. ROC curve of average gymnastic activity identification accuracy.

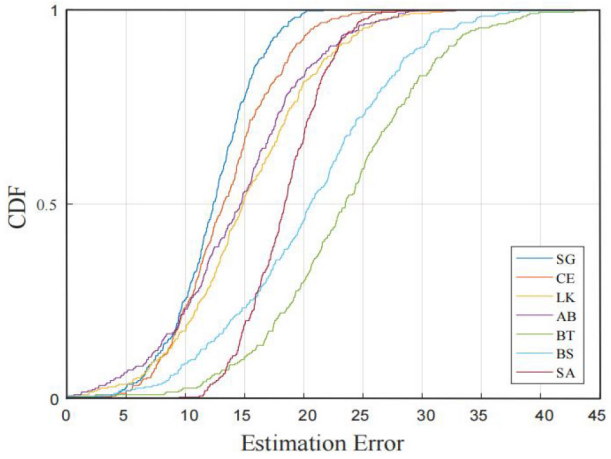


Fig. 13. CDF of estimation errors for the seven activities.

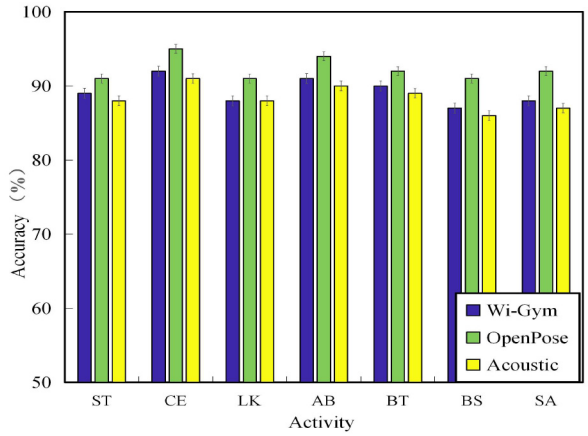


Fig. 12. Evaluation accuracy of different gymnastics activities.

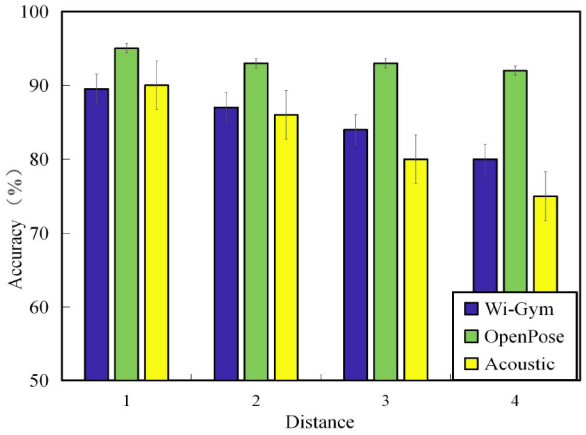


Fig. 14. Impact of the exerciser's position on evaluation accuracy.

C. Gymnastics Activity Evaluation Performance

1) Impact of Different Activities: The impact of the activity type on the evaluation accuracy is studied. The accuracy is defined as the ratio of the number of correctly evaluated instances (doing gymnastics and judged by the trainer, with a score is higher than 80) to the total number of instances. The same activity is performed and monitored using three kinds of sensing methods. The results are presented in Figs. 12 and 13. For the seven basic activities, the average activity evaluation accuracy of Wi-Gym ranges from 87% (BS) to 92% (CE); the average accuracy of the video-based approach ranges from 91% (ST) to 95% (CE); and average accuracy of the acoustic-based method ranges from 86% (BS) to 91% (CE). These results demonstrate the capability of Wi-Gym to handle various types of activities. The video-based scheme performs the best because the Gabor filter is originally designed for image processing, and the transformed RF signals are usually distorted. The acoustic-based scheme and Wi-Gym have similar performance, both being sensitive to the environmental interference. The acoustic and Wi-Fi-based schemes perform the best for the CE activity and the worst for the BS activity. With the current setting (i.e., one transmitter and one receiver), the Wi-Fi or acoustic measurements mainly reflect the activity information perpendicular to the LoS path. Therefore, the activities involving the movements perpendicular to the LoS path are easier to detect. On the contrary, the activities involving the movements parallel to the LoS path are more difficult to identify.

2) Impact of the Exerciser's Position: To study the impact of the exerciser's position on evaluation accuracy, the exercisers perform the gymnastic exercises at different positions during this experiment. When Wi-Gym and the acoustic-based system are deployed, the distance from the exerciser to the Line-of-Sight (LoS) path between the transmitter and receiver is varied from 1 to 4 m. As for the video-based system, the distance between the exerciser and the camera is varied from 1 to 4 m. Each subject performs the gymnastics exercises at different positions, and the evaluation results are presented in Fig. 14. When the distance from the subject to the LoS is increased from 1 to 3m, Wi-Gym can always maintain a pretty good performance. However, when the distance reaches 4 m, the evaluation accuracy drops significantly. The performance of the video-based system does not decrease significantly with the increased distance because video data are not so sensitive to distance. The performance of the acoustic-based system is inferior to Wi-Gym, since acoustic signals attenuate more seriously than Wi-Fi signals when the distance gets larger. This

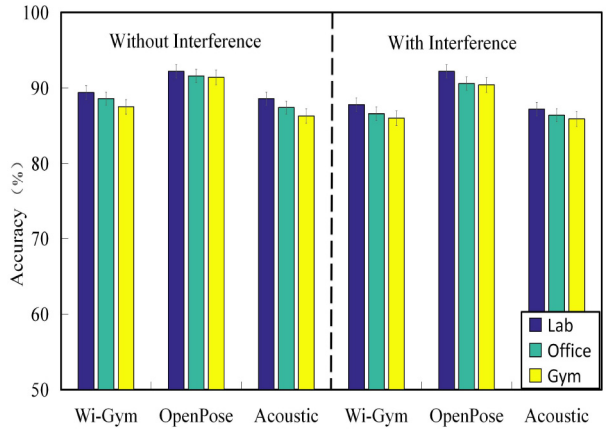


Fig. 15. Average gymnastics activity assessment accuracy under layout changes.

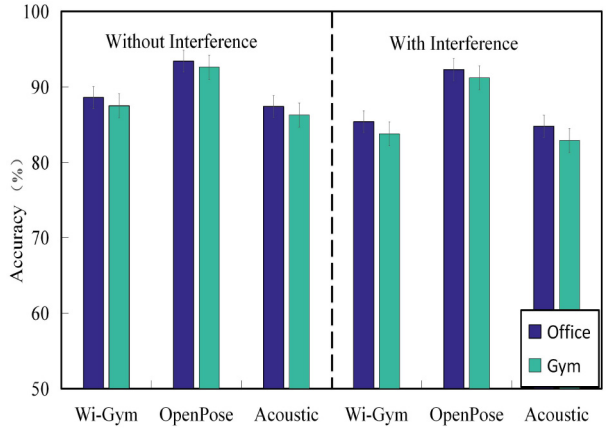


Fig. 16. Average gymnastics activity assessment accuracy under environmental changes.

makes fine-grained activity evaluation more difficult. The maximum area for Wi-Gym to work well is 3 m × 3 m. The Wi-Fi transmit power can be increased or antennas with higher gain can be used to improve the sensing range of Wi-Gym. The experiments validate the robustness to distance changes of the three prototype systems within their sensing ranges.

3) Impact of the Environment: In order to evaluate Wi-Gym's adaptability to environment, tests are carried out in various environments. We experiment with the case of entire environment changes and the case of the same environment but with changed layouts. Wi-Gym is trained in the original scenario and tested in different, changed scenarios.

The results before and after the environmental change are shown in Figs. 15 and 16. In Fig. 15, the environmental changes lead to a similar trend in the laboratory, office, and gym results. It is obvious that the video-based system is the most accurate and least affected by environmental changes. As the layout is changed, the signal reflection path variation degrades the performance of Wi-Gym. The average accuracy of laboratory, office, and gym experiments decreases from 89.5%, 88.6%, and 87.5% to 87.8%, 86.8%, and 86%, respectively. The results indicate that furniture relocation has little impact on evaluation accuracy. The acoustic-based system's

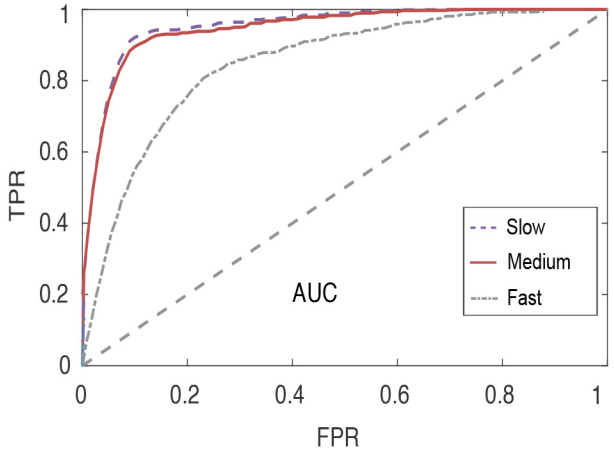


Fig. 17. Impact of the diversity in the speed of performing gymnastics activities.

performance is similar to that of the Wi-Fi-based system. When the environment is changed, the assessment accuracies all slightly decrease. This verifies that the domain adaptation methodology employed in Wi-Gym effectively mitigates the performance degradation caused by environmental changes.

In Fig. 16, the average accuracy of Wi-Gym in office and gym declines from 89.5% and 88.5% to 86.4% and 83.8%, respectively. The experimental results show that even tested in a new environment, the evaluation accuracy will only drop slightly. Obviously, the domain adaptation methodology plays a significant role on maintaining the gymnastic activity evaluation accuracy. The domain adaptation methodology effectively addresses the source and target domain imbalance issue. The similar performance of the three sensing methods also verifies their robustness to environmental changes.

4) Impact of Speed Diversity: According to [40], a duration of each activity is usually in the range from 1 to 3 s. In reality, different people have different ways of performing gymnastic activities, usually at different paces or speeds. The impact of the gymnastic activity speed is studied in this experiment, considering slow (duration is above (t + 0.3) s), medium (duration is within (t − 0.3) s, (t + 0.3) s), and fast (duration is below (t − 0.3) s) cases. Fig. 17 depicts the accuracy results for different speed levels. With certain sampling rate, when the speed is the lowest, the performance is the best. With the increase of speed, Wi-Gym's performance degrades. The faster the speed, the less complete the captured signals. In addition, doing an activity at a faster speed, the exerciser is more likely to perform it in a hurry with larger distortions. The accuracy degradation in Fig. 17 validates this fact. On the contrary, when an activity is carried out at the normal speed or slow speed, the accuracy is usually much better.

5) Impact of Clothing: Wi-Gym is first trained with data samples collected when a subject wears different clothes. Then, the accuracy of Wi-Gym trained with different clothes will be tested. Fig. 18 plots the accuracy of Wi-Gym with various seasonal clothes. It is obvious that clothes have certain impacts on the evaluation accuracy performance. When trained and tested with the clothes in summer (winter), the test accuracy of winter (summer) is the worst. With tested

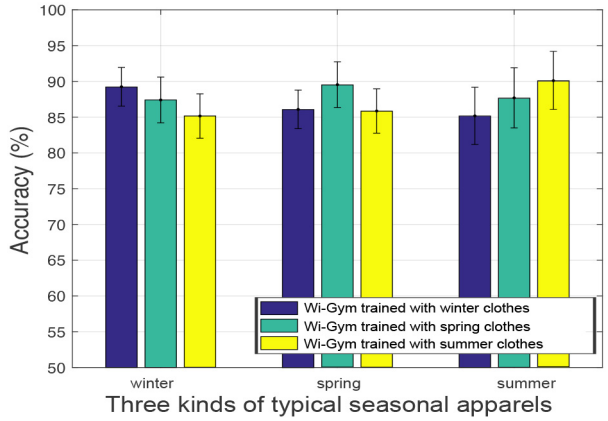


Fig. 18. Impact of the exerciser's clothing on the accuracy of Wi-Gym.

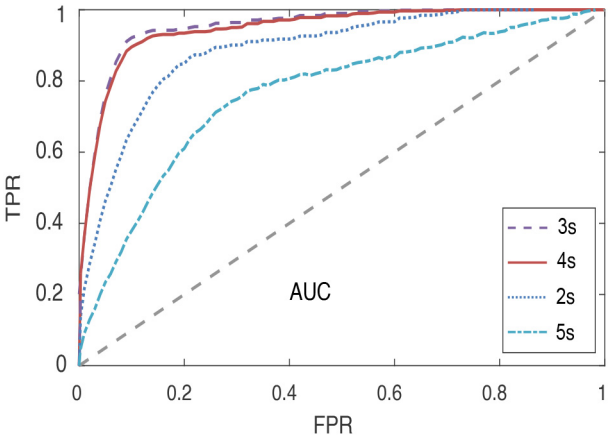


Fig. 19. Impact of window size.

with spring clothes, the accuracy will be medium, no matter trained with winter clothes or summer clothes. The reason behind this is that spring clothes is neither too thin nor too thick, lying between summer and winter clothes. In general, despite that Wi-Gym is never trained on the samples with the other two seasonal clothes, it can still maintain an accuracy above 85% in all the tests.

6) Impact of Window Size: Usually the changes in the activity speed are not beneficial to activity evaluation. In this experiment, different sliding window sizes from 2 to 5 s are evaluated and compared, and the results are presented in Fig. 19. When the window size is set to 3 s, Wi-Gym performs the best. When the window size is set to 5 s, Wi-Gym performs the worst. This is because the influence of noise on activity evaluation is significantly amplified when the sliding window becomes large.

7) Impact of Lighting Condition and Blockage: To assess the influence of the lighting condition on activity evaluation, several experiments are conducted in the environment under various lighting conditions. In the experiments, 50 subjects, including 25 males and 25 females, are tested in an office, a laboratory, and a gymnasium. All the experiments are illuminated with nature lights through windows, while the electric lights are all turned off. Each subject stands in front of the LoS path and performs the seven basic activities. An Xbox Kinect 2.0 device is located at the center of the LoS path at a height of 80 cm. The Wi-Fi and video measurements are collected simultaneously. Both classification models are trained with the data collected in normal lighting conditions and tested with the data collected in dimmed lighting conditions. The ground-truth scores are the average scores given by several gymnastic experts. Every subject repeats each activity 100 times in each scenario. The average scores computed by the sensing systems are compared with the ground truth.

TABLE II
ACCURACY UNDER NORMAL LIGHTING AND DIMMED LIGHTING CONDITIONS

	Normal Light			Dimmed Light		
	Laboratory	Office	Gymnasium	Laboratory	Office	Gymnasium
Wi-Gym	89.4%	88.6%	87.5%	89.2%	88.3%	86.4%
Video	92.2%	91.6%	91.4%	83.8%	82.4%	80.7%

TABLE III
ACCURACY WITH OR WITHOUT OBSTACLE

	With Obstacle			Without Obstacle		
	Laboratory	Office	Gymnasium	Laboratory	Office	Gymnasium
Wi-Gym	89.4%	88.6%	87.5%	88.2%	87.3%	85.4%
Video	92.2%	91.6%	91.4%	84.6%	83.5%	81.2%

We conduct two types of experiments: 1) under different lighting conditions and 2) under different blockage conditions. To evaluate the impact of lighting, experiments are conducted at 9:30 (i.e., under normal lighting conditions) and 21:30 (under dimmed light). The accuracy results are presented in Table II. As expected, the video-based method is quite sensitive to the lighting condition. Its performance does not compete with Wi-Fi-based sensing when the lighting condition is poor. However, since the SVM classifier is enhanced with domain adaptation, the video-based method still achieves a satisfactory accuracy (i.e., over 80%) in all the three scenarios when there is a change in the lighting condition.

To evaluate the impact of obstacles, a chair is placed between the LoS and the subject during the next set of experiments. The classifiers are trained with the data collected without the chair and tested with the data with the chair. As shown in Table III, Wi-Gym outperforms the video-based method when an obstacle is introduced. Both schemes still perform well with the worst case accuracy higher than 81.2% even under blockage.

8) Impact of Transceiver Height: To explore the influence of transceiver height on activity evaluation, the transceiver height is varied from 1 to 1.5 m in this experiment. Fifty subjects perform the seven basic activities one by one at a location that is 1 m away from the LoS path, facing the LoS path. The corresponding evaluation results are plotted in Fig. 20. The experimental results show that when the height of transceiver is changed, the evaluation accuracy does not change significantly. In summary, Wi-Gym is resilient to transceiver height changes.

9) System Complexity Analysis: In order to understand the real-time performance of Wi-Gym, the computational complexity of each processing procedure is analyzed and their

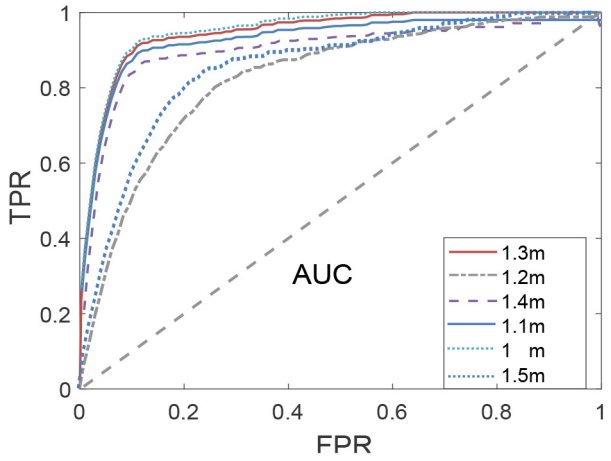


Fig. 20. Impact of the height of the Wi-Fi transmitter/receiver.

TABLE IV
COMPUTATIONAL COMPLEXITY AND TIME CONSUMPTION OF PROCESSING 10-S CSI DATA

Module	Complexity	CPU Time (s)
Noise Removal	$O(\min(L^3, N_s^3 N_c^3)))$	0.8863
Movement Segmentation	$O(L)$	0.0502
Feature Extraction	$O(Llog_2 L)$	0.2034
Classification	$O(D^3 + ND^2)$	0.0005
Normativity Evaluation	$O(ND^2)$	0.0005
Total	$O(\min(L^3, N_s^3 N_c^3)))$	1.1409

time consumptions are calculated. Wi-Gym is evaluated with two T-series Lenovo laptops. Each laptop is equipped with an Intel 5300 network interface card and Ubuntu desktop 11.04 OS. The time consumption and computational complexity of processing 10 s of CSI data are provided in Table IV. The length of the CSI signal is denoted by L, $N_c = 3$ represents the number of CSI streams, $N_s = 30$ represents the number of subcarriers, and $N_v = 10$ denotes the first ten PCA components. N and D represent the number of training instances and the dimension of the feature, respectively. The main computational cost is incurred on data calibration (which involves PCA processing). In the process of data calibration, the data size is reduced to one ninth of the original data. The testing time is only 1.05 s for evaluating one activity. The highly efficient testing process enables Wi-Gym to evaluate activities continuously in real time. In summary, the complexity analysis implies that the trained Wi-Gym can ensure good real-time performance.

V. FURTHER DISCUSSION

The experimental study has demonstrated the robustness and effectiveness of Wi-Gym. Compared with alternative approaches, the vision-based method works best when the lighting condition is good. When the lighting condition is not good, Wi-Gym works better than the vision-based method. In addition, if there is an obstacle between the exerciser and LoS, the vision-based method's performance gets worse and cannot compete with Wi-Gym. The vision-based method usually raises the privacy concern, making it unsuitable for certain applications. The acoustic-based method usually achieves a similar performance as Wi-Gym until the distance between the exerciser and the LoS path is larger than 1 m. The acoustic-based method is very sensitive to the distance. When the distance gets larger, its performance degrades and becomes worth than Wi-Gym. In addition, Wi-Gym is not a simple gymnastic activity recognition system, as the vision or acoustic-based prior works. It has the unique capability of gymnastic activity scoring. There is hardly any gymnastic activity scoring system available in the literature, even for the vision-based system. Wi-Gym is an effective first-of-its-kind gymnastic activity assessing system utilizing commercial Wi-Fi. It has the advantage of environment dependence mitigation since its SVM classifier has been enhanced with domain adaptation. Therefore, when there are environmental changes, Wi-Gym can adapt to the changes and still maintain a good performance.

However, the current Wi-Gym can be further improved from the following two aspects.

1) *Multiperson Scenario:* As analyzed in [41], the movement interference generated from surrounding people can be ignored if they are moving outside the sensing region. Wi-Gym is focused on the evaluation of one subject's activity in a typical indoor scenario. Multiple stationary subjects in the test area have no impact on Wi-Gym. When there are some moving subjects around and they are in the system's sensing range, the received signals reflected from multiple subjects will be mixed together at the Wi-Fi receiver. Due to the lack of CSI spatial information, it will be challenging to derive each exerciser's movement-induced signals. The successive cancelation-based techniques relying on the separation of distance and azimuth of multiple subjects may not be suitable for Wi-Gym. How to extract the activity-induced signals of each exerciser will be a good problem for further study.
2) *Scalability*: Though Wi-Gym can be customized to a wide range of applications, such as smart home or smart sports, it still has some scalability issues. Furthermore, when the exerciser is far away from the Wi-Fi devices. Wi-Gym's performance will become worse. It is necessary to deploy more Wi-Fi devices to balance the interference and sensing range and maintain a satisfactory evaluation accuracy. Therefore, optimizing the Wi-Fi deployment, improving the scalability of the system, restraining the interference, and guaranteeing a satisfactory performance should be studied in future work.

VI. CONCLUSION

We proposed Wi-Gym, a commodity Wi-Fi-based framework for evaluating the quality of gymnastics activities. A strategy for gymnastics activity assessment was proposed and implemented by utilizing DTW to compare an exerciser activity-induced CSI dynamics with a trainer activity-induced CSI dynamics. FIS was incorporated to provide the evaluation score based on the DTW results. Wi-Gym also leveraged domain adaptation to mitigate the impact of changes in the

deployment environment. The extensive experiments validated the effectiveness and robustness of Wi-Gym. Extending Wi-Gym to the case of multiple exercisers and enhancing its scalability will be interesting problems for future work.

REFERENCES

[1] M. P. Diaz-Pereira, I. Gomez-Conde, M. Escalona, and D. N. Olivieri, "Automatic recognition and scoring of olympic rhythmic gymnastic movements," *Human Movement Sci.*, vol. 34, no. 1, pp. 63–80, Apr. 2014.

[2] L. A. Sanchez-Perez, L. P. Sanchez-Fernandez, A. Shaout, J. M. Martinez-Hernandez, and M. J. Alvarez-Noriega, "Rest tremor quantification based on fuzzy inference systems and wearable sensors," *Int. J. Med. Inform.*, vol. 114, no. 3, pp. 6–17, Mar. 2018.

[3] R. Niewiadomski, K. Kolykhalova, S. Piana, P. Alborno, G. Volpe, and A. Camurri, "Analysis of movement quality in full-body physical activities," *ACM Trans. Interact. Intell. Syst.*, vol. 9, no. 1, pp. 1–20, Mar. 2019.

[4] H. Ghasemzadeh and R. Jafari, "Coordination analysis of human movements with body sensor networks: A signal processing model to evaluate baseball swings," *IEEE Sensors J.*, vol. 11, no. 3, pp. 603–610, Mar. 2011.

[5] H.-T. Cheng, F.-T. Sun, M. Griss, P. Davis, J. Li, and D. You, "NuActiv: Recognizing unseen new activities using semantic attribute-based learning," in *Proc. ACM MobiSys*, Taipei, Taiwan, Jun. 2013, pp. 361–374.

[6] M. Chen, Y. Li, X. Luo, W. Wang, L. Wang, and W. Zhao, "A novel human activity recognition scheme for smart health using multilayer extreme learning machine," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1410–1418, Apr. 2019.

[7] D. Tao, Y. Wen, and R. Hong, "Multicolumn bidirectional long short-term memory for mobile devices-based human activity recognition," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 1124–1134, Dec. 2016.

[8] H. Huang, X. Li, S. Liu, S. Hu, and Y. Sun, "triboMotion: A self-powered Triboelectric motion sensor in wearable Internet of Things for human activity recognition and energy harvesting," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4441–4453, Dec. 2018.

[9] B. Fang, N. D. Lane, M. Zhang, A. Boran, and F. Kawsar, "BodyScan: Enabling radio-based sensing on wearable devices for contactless activity and vital sign monitoring," in *Proc. ACM MobiSys*, Singapore, Jun. 2016, pp. 97–110.

[10] A. Allahdadi and R. Morla, "Anomaly detection and modeling in 802.11 wireless networks," *J. Netw. Syst. Manage.*, vol. 27, no. 2, pp. 3–38, Jan. 2019.

[11] A. Benzerbadj, B. Kechar, A. Bounceur, and M. Hammoudeh, "Surveillance of sensitive fenced areas using duty-cycled wireless sensor networks with asymmetrical links," *J. Netw. Comput. Appl.*, vol. 112, no. 3, pp. 41–52, Jun. 2018.

[12] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. ACM MobiCom*, Paris, France, Sep. 2015, pp. 65–76.

[13] Y. Wang, K. Wu, and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 581–594, Feb. 2017.

[14] H. Wang, D. Zhang, Y. Wang, J. Ma, and S. Li, "RT-fall: A real-time and contactless fall detection system with commodity WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 511–526, Feb. 2017.

[15] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, p. 121, Sep. 2019.

[16] J. Yang, H. Zou, H. Jiang, and L. Xie, "Device-free occupant activity sensing using WiFi-enabled IoT devices for smart homes," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3991–4002, Oct. 2018.

[17] P. Khan, B. S. K. Reddy, A. Pandey, S. Kumar, and M. Youssef, "Differential channel-state-information-based human activity recognition in IoT networks," *IEEE Internet Things J.*, vol. 7, no. 11, pp. 11290–11302, Nov. 2020.

[18] W. Aigner, M. Tomitsch, M. Stroe, and R. Rzepa, "Be a judge! Wearable wireless motion sensors for audience participation," in *Proc. CHI*, Vienna, Austria, Apr. 2004, pp. 1617–1621.

[19] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proc. ACM MobiCom*, Maui, HI, USA, Sep. 2014, pp. 617–628.

[20] Y. Zheng *et al.*, "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proc. ACM MobiSys*, Seoul, South Korea, Jun. 2019, pp. 313–325.

[21] R. Ayyalasomayajula *et al.*, "Deep learning based wireless localization for indoor navigation," in *Proc. ACM MobiCom*, London, U.K., Apr. 2020, pp. 1–14.

[22] Y. Cao *et al.*, "Contactless body movement recognition during sleep via WiFi signals," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2028–2037, Mar. 2020.

[23] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "WiDir: Walking direction estimation using wireless signals," in *Proc. ACM UbiComp*, Heidelberg, Germany, Sep. 2016, pp. 351–362.

[24] K. Qian, "Passive human sensing with COTS Wi-Fi," in *Proc. ACM MobiSys*, Munich, Germany, Jun. 2018, pp. 3–4.

[25] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single Wi-Fi link," in *Proc. ACM MobiSys*, Munich, Germany, Jun. 2018, pp. 350–361.

[26] Y. Gu, J. Zhan, Y. Ji, J. Li, F. Ren, and S. Gao, "MoSense: An RF-based motion detection system via off-the-shelf WiFi devices," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2326–2341, Dec. 2017.

[27] W. Jiang *et al.*, "Towards 3D human pose construction using WiFi," in *Proc. ACM MobiCom*, London, U.K., Apr. 2020, pp. 1–14.

[28] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using WiFi signals," in *Proc. ACM MobiCom*, Paris, France, Sep. 2015, pp. 90–102.

[29] J. C. P. Chan, H. Leung, J. K. T. Tang, and T. Komura, "A virtual reality dance training system using motion capture technology," *IEEE Trans. Learn. Technol.*, vol. 4, no. 2, pp. 187–195, Apr.–Jun. 2011.

[30] J. R. Movellan, "Tutorial on Gabor filters," *Open Source Doc.*, vol. 49, no. 2, pp. 1–23, Feb. 2002.

[31] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu, "Boosting for transfer learning," in *Proc. ICML*, Corvallis, OR, USA, Jun. 2007, pp. 193–200.

[32] W. Li, Z. Xu, D. Xu, D. Dai, and L. Van Gool, "Domain generalization and adaptation using low rank exemplar SVMs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1114–1127, May 2018.

[33] Y. He, Y. Tian, J. Tang, and Y. Ma, "Unsupervised domain adaptation using exemplar-SVMs with adaptation regularization," *Complex.*, vol. 2018, Jan. 2018, Art. no. 8425821.

[34] T. M. Williams, "Practical methods of optimization," *J. Oper. Res. Soc.*, vol. 32, p. 417, May 1981.

[35] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Disc.*, vol. 2, no. 2, pp. 121–167, Feb. 1998.

[36] T. Knebel, S. Hochreiter, and K. Obermayer, "An SMO algorithm for the potential support vector machine," *Neural Comput.*, vol. 20, no. 1, pp. 271–287, Jan. 2008.

[37] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man-Mach. Stud.*, vol. 7, no. 1, pp. 1–13, Jan. 1975.

[38] M. A. Masoum and E. F. Fuchs, "Index-power quality in power systems and electrical machines," in *Proc. Power Qual. Power Syst. Electr. Mach.*, 2015, pp. 1105–1123.

[39] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, Jul. 2021.

[40] Z. Jun, "A probe into the development tendency of broadcast gymnastics in the 21st century," *J. Hubei Sports Sci.*, vol. 20, pp. 271–287, Jan. 2003.

[41] X. Guo, J. Liu, C. Shi, H. Liu, Y. Chen, and M. C. Chuah, "Device-free personalized fitness assistant using WiFi," in *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, p. 165, Dec. 2018.

Lei Zhang (Member, IEEE) received the Ph.D. degree in computer science from Auburn University, Auburn, AL, USA, in 2008.

From 2008 to 2011, she worked as an Assistant Professor with the Computer Science Department, Frostburg State University, Frostburg, MD, USA. She is currently an Associate Professor with the College of Intelligence and Computing, Tianjin University, Tianjin, China. Her research interests include mobile computing and data mining.

Dr. Zhang is a member of ACM.

Wenyuan Huang received the B.E. degree from Hainan University, Hainan, China, in 2020. He is currently pursuing the master's degree with Tianjin University, Tianjin, China.

His research interests include wireless sensing and machine learning.

Xiaoxia Jia received the B.E. degree from Taiyuan Normal University, Jinzhong, Shanxi, China, in 2017. She is currently pursuing the master's degree with Tianjin University, Tianjin, China.

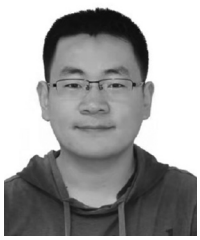
Her research interests include wireless communication and data mining.

Xiaojie Fan received the B.E. degree from Tiangong University, Tianjin, China, in 2019. She is currently pursuing the master's degree with Tianjin University, Tianjin.

Her research interests include wireless communication and data mining.

Xiaochen Fan received the B.E. degree from Beijing Institute of Technology, Beijing, China, in 2013, and the Ph.D. degree from the University of Technology Sydney, Ultimo, NSW, Australia, in 2021.

He is a Postdoctoral Researcher with the Department of Electronic Engineering, Tsinghua University, Beijing, China. He also works as an Assistant Researcher with the Consulting Center for Strategic Assessment, Academy of Military Science, Beijing, China. His research interests include mobile computing, Internet of Things, spatio-temporal data mining, and deep learning.



Liangyi Gong (Member, IEEE) received the B.S. and Ph.D. degrees from the School of Computer Science and Technology, Harbin Engineering University, Harbin, China, in 2010 and 2016, respectively.

He is a Postdoctoral Researcher with the School of Software and BNRist, Tsinghua University, Beijing, China. His research interests are in the areas of network security and mobile computing.

Wenyuan Tao received the Ph.D. degree from Tianjin University, Tianjin, China, in 2002.

He is a Professor with the College of Intelligence and Computing, Tianjin University. His research interests include big data analysis and machine learning.

Shiwen Mao (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Polytechnic University, Brooklyn, NY, USA, in 2004.

He is currently a Professor and an Earle C. Williams Eminent Scholar of Electrical and Computer Engineering with Auburn University, Auburn, AL, USA. His research interests include wireless networks, multimedia communications, and smart grid.

Prof. Mao received the IEEE ComSoc TC-CSR Distinguished Technical Achievement Award in 2019, the Auburn University Creative Research and Scholarship Award in 2018, and the NSF CAREER Award in 2010. He is a co-recipient of the 2021 Best Paper Award of Elsevier/KeAi *Digital Communications and Networks* Journal, the 2021 IEEE INTERNET OF THINGS JOURNAL Best Paper Award, the 2021 IEEE Communications Society Outstanding Paper Award, the IEEE Vehicular Technology Society 2020 Jack Neubauer Memorial Award, the IEEE ComSoc MMTC 2018 Best Journal Paper Award and the 2017 Best Conference Paper Award, the Best Demo Award of IEEE INFOCOM 2022 and IEEE SECON 2017, the Best Paper Awards of IEEE ICC 2022 and 2013, IEEE GLOBECOM 2019, 2016, and 2015, and IEEE WCNC 2015, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems. He is a Distinguished Lecturer of IEEE Communications Society and IEEE Council of RFID, and was a Distinguished Lecturer from 2014 to 2018 and a Distinguished Speaker from 2018 to 2021 of IEEE Vehicular Technology Society. He is on the Editorial Board of IEEE/CIC CHINA COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, ACM GetMobile, IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE MULTIMEDIA, IEEE NETWORK, and IEEE NETWORKING LETTERS.